Methods and kits for diagnosing and treating B-Cell Chronic Lymphocytic Leukemia (B-CLL)

All patent and non-patent references cited in the present application, are hereby incorporated by reference in their entirety.

Field of invention

The present invention relates to methods and kits for detecting several polynucleotide sequence found to be indicative of a poor prognosis of B-CLL. All the polynucleotides are transcribed from a region on human chromosome 12p21-22. Most of the polynucleotides do not encode larger polypeptides, but may encode small peptides, they may function as RNAs. Four polynucleotides encode a novel protein, which in one preferred embodiment can be used as a cytokine, preferably as an interleukin. Furthermore the invention relates to methods and compositions for treating B-CLL in particular poor prognosis B-CLL.

Background of invention

B-CLL is the most common form of leukaemia in Denmark, with more than 250 new cases diagnosed every year. The disease results in accumulation of CD19+CD5+CD23+ lymphocytes in the blood, bone marrow and organs of the patients. B-CLL cells are long-lived, slowly dividing and locked in the G_1 phase of the cell cycle. At this time it is unknown how or why B-CLL occurs and no cure is known for B-CLL. The application of more aggressive treatment strategies has been hampered by the inability to identify reproducible and reliable prognostic predictors in patients with poor outcome in this disease. In many patients the diagnosis does not affect morbidity or mortality. Other patients suffer from an incurable cancer that inevitably results in death, regardless of treatment. Until recently this latter group of patients could not be identified at the time of diagnosis. Recently, two studies established the mutational status of immunoglobulin variable region of the heavy chain (Ig V_H) genes in B-CLL as independent prognostic markers, within each clinical stage (Damle, et al. & Hamblin, et al.). Patients without somatic hypermutation show much shorter survival than patients with somatic hypermutation. FISH-studies of cytogenetic aberrations in B-CLL established specific abnormalities on chromosomes 11 (ATM), 12 (?), 13 (Leu-1 and-2) and 17 (p53) as independent prognostic markers, within each clinical stage (Dohner, et al.). Very recent studies have demonstrated that independent risk prediction, using a combined analysis of Ig V_H gene mutational analysis and cytogenetics, can identify subgroups of B-CLL with median survivals ranging from less than 2.5 years to more than 15 years (Krober, et al., Lin, et al., & Oscier, et al.) (see Figure 1). Since the process of characterising the Ig VH gene mutational status of an individual patient is cumbersome, it is desirable to provide easier tests based on diagnostic markers for use in the differential diagnosis of such cancer patients.

Summary of invention

It is an object of preferred embodiments of the present invention to provide differentially expressed transcription products, which can be used as prognostic markers of disease and give information about the differences in etiology between different groups of B-CLL patients. These differentially expressed transcription products are genetic markers that can be used in an easy assay to distinguish between subgroups of B-CLL patients and especially identify B-CLL patients with a poor prognosis.

This method for diagnosing a subtype of B-cell chronic lymphocytic leukaemia (B-CLL) comprises the steps of determining the presence or absence of at least one expression product such as a transcriptional product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in a biological sample isolated from a individual. As evidenced by the appended examples, the present inventors have determined that the expression products of this invention are present in one subtype of B-CLL having poor prognosis and thus of great diagnostic value and independent prognostic value. Equally important, an expression product comprising a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 has not been found in any of the other tissue types tested (see e.g. Figure 8).

The vast majority of patients which show expression of the AMB-1 gene in form of at least one of the expression products selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 show unmutated Ig V(H) genes which is consistent with poor prognosis B-CLL. The presence of an expression product of the AMB-1 gene can be determined easily using standard laboratory procedures and equipment. Therefore the diagnostic method provided by the present inventors provides an easy method of diagnosis as compared to the determination of the mutation status of Ig V(H) genes and can furthermore give additional information about the prognosis.

Accordingly, a further object of preferred embodiments of the present invention is a method for determining the progress of B-CLL comprising determining the amount of at least one expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in a biological sample isolated from an individual. The method may be used e.g. for determining the efficiency of a treatment, i.e. to see whether the amount of the expression product decreases or increases in response to a curative treatment.

The expression products of the present invention are all transcripts of SEQ ID NO:1 and/or SEQ ID No:5, the gene of the present invention called AMB-1 which also encodes a novel polypeptide (SEQ ID NO:3).

A further object of preferred embodiments of the present invention is to provide a cure and/or treatment of patients with B-CLL, in particular of patients with poor prognosis B-CLL such as the sub-type of B-CLL which is characterised by the presence of an expression product of the present invention.

The method for treating B-CLL comprises administering to an individual with a B-CLL diagnosis a compound capable of decreasing or inhibiting the formation of an expression product of SEQ ID NO:1 and/or SEQ ID NO:5. This expression product preferably comprises a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18. The present inventors believe that the presence of at least one of said expression products is an etiological factor in B-CLL and that the disease can be treated or cured by inhibiting the expression of at least one of such products and/or by inhibiting the effect of such product by e.g. rendering it inactive.

A further preferred object of embodiments of the present invention is to destroy or to eliminate the transcription of at least one expression product comprising at least one nucleotide sequence selected form the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 for the treatment of cancer, such as a poor prognosis sub-type of B-CLL.

The above destruction or elimination is obtained by applying polynucleotides or oligonucleotides in the form of small interfering RNA molecules (siRNA), antisense molecules or ribozymes.

In one aspect the invention relates to a gene therapy vector capable of inhibiting or decreasing the formation of an expression product of SEQ ID NO:1 and/or SEQ ID NO:5, said gene therapy vector preferably encoding a specific siRNA molecule, a specific antisense molecule or a specific ribozyme being capable of decreasing or inhibiting the formation of an expression product of SEQ ID NO:1 and/or SEQ ID NO:5. This gene therapy vector can be used for treating B-CLL based on the finding that the AMB-1 gene encoded by SEQ ID No:1 and/or SEQ ID No:5 is an etiological factor in B-CLL.

Both SEQ ID No 1 which is a 20,000 nucleotide long sequence and SEQ ID No 5 which is a 80,000 nucleotide long sequence provides several transcriptional products in B-CLL cells in patients with poor prognosis B-CLL. Some of the transcriptional products e.g. SEQ ID No 2 and SEQ ID No 4 consists of two exons (SEQ ID No: 15 and SEQ ID No: 16) separated by the same intron. Both mRNA sequences encode an open reading frame (SEQ ID No: 17) encoding a 121 amino acid peptide (SEQ ID No 3).

Accordingly, yet another object of preferred embodiments of the present invention relates to a novel class of polypeptides. These may be described as a group of isolated polypeptides or proteins comprising or essentially consisting of the amino acid sequence of SEQ ID No. 3, or a fragment thereof, or a polypeptide functionally equivalent to SEQ ID

No. 3, or a fragment thereof, wherein said fragment or functionally equivalent polypeptide has at least 60% sequence identity with the polypeptide of SEQ ID No 3. The polypeptides of the present invention may have interleukin or cytokine activity.

In a still further aspect the invention relates to an isolated polynucleotide selected from the group consisting of:

i) a polynucleotide comprising nucleotides of SEQ ID No 5,
ii) a polynucleotide encoding a polypeptide having the amino acid sequence of SEQ ID No 3,
iii) a polynucleotide, the complementary strand of which hybridises, under stringent conditions, with a polynucleotide as defined in any of i) and ii).
iv) a polynucleotide which is degenerate to the polynucleotide of iii), and
v) the complementary strand of any such polynucleotide.

The polypeptides encoded by the polynucleotides may furthermore

a) have at least 60 % sequence identity with the amino acid sequence of SEQ ID No 3 and have interleukin or cytokine activity,
b) be recognised by an antibody, or a binding fragment thereof, which is capable of recognising an epitope, wherein said epitope is comprised within a polypeptide having the amino acid sequence of SEQ ID No 3; and/or
c) be competing with a polypeptide having the amino acid sequence as shown in SEQ ID No 3 for binding to at least one predetermined binding partner such as a cytokine receptor.

One further therapeutic application of the present invention is a method of vaccination against B-CLL said method comprising immunising an individual against a translational product of SEQ ID No:1 and/or SEQ ID No:5. By stimulating the immune system of an individual to produce antibodies against the translational product the individual can become immune towards B-CLL and/or the method can be used as part of therapy. The state of the art describes various ways of immunising an individual against a particular protein.

Finally, the invention provides a method for determining an increased or decreased predisposition for B-CLL comprising determining in a biological sample from an individual a germline alteration in a target nucleic acid sequence comprising 150,000 nucleotides, said target nucleic acid sequence comprising at least 10 nucleotides of SEQ ID No:1 and/or SEQ ID No:5. This aspect is based on the finding of the importance of the expression product of SEQ ID No:1 and/or SEQ ID No:5, and the absence of any detectable expression product of SEQ ID No:1 and/or SEQ ID No:5 in healthy tissue and in patients with good prognosis B-CLL. It is highly likely that the difference is caused by a germline alteration. A germline alteration can be targeted by gene therapy methods and by the methods provided in the present invention.

Description of Drawings

Figure 1: Overall survival of B-CLL patients by genotype (all stages) The prognostic significance of V_H homology and cytogenetic aberrations is independent of clinical stage (from Kröber et al., 2002 (4)).

Figure 2 (a-d): Survival curves for survival or time to progression based on *AMB-1* expression or IgV_H mutational status respectively. Patients are newly diagnosed, untreated B-CLL patients (n=34).

Figure 3: Northern blotting demonstrating the expression level of AMB-1 in various tissue samples using as a probe an 875 base pair fragment of Exon 3/Seq ID No:16. UPN 1, 4 and 7 are unmutated B-CLL patients, UPN 19, 9, 10, 13 and 21 are mutated B-CLL patients. Included are RNA samples from normal Colon, Spleen, Bone Marrow and PBL (peripheral blood lymphocytes) and RNA from the Ramos and Granta cell lines. Equal loading of lanes was confirmed by re-probing with an actin probe (results not shown).

Figure 4. Alignment of AMB1 with IL4 based on structurel similarity. IL4 is called d1iara in the alignment. The additional lines indicate the structural similarity.

Figure 5. A 3D search, where the peptide sequence has been searched for similarity to known protein or peptide 3D-structures.

Figure 6. Predicted 3-D structure of AMB-1 compared to the known 3-D structure of human IL4. Prediction is performed using SEQ ID No:3 and the method described in: Enhanced Genome Annotation using Structural Profiles in the Program 3D-PSSM. Kelley LA, MacCallum RM & Sternberg MJE (2000). J. Mol. Biol. 299(2), 499-520.

Figure 7. Alignment of the AMB1 peptide sequence with the sequences of IL4, IL3, IL13 and GM-CSF, based on their structures.

Figure 8. A table showing the tissue types on the MTE array used for dot blotting of AMB-1 to check for expression in other tissue types.

Figure 9. Schematic representation of the transcriptional products of the present invention compared to the genomic AMB1 sequence (1 & 5). 2 is mRNA short form (SEQ ID No 2). 4 is mRNA long form (SEQ ID No 4). 6 (SEQ ID No:6), 7 (SEQ ID No:7), 8 (SEQ ID No:8), 9 (SEQ ID No:9), 10 (SEQ ID No:10) and 11 (SEQ ID No:11) are alternative transcription products all comprising at least one nucleotide sequence selected form the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18.

Figure 10. RT-PCR demonstrating the expression of AMB-1 in B-CLL patients by RT-PCR. UPN1-UPN8 are unmutated patients, UPN9-UPN16 are mutated patients.

Detailed description of the invention

The present invention in particular relates to polynucleotide sequences found to be indicative of a poor prognisis of B-CLL and new methods and compositions for treating B-CLL in particular poor prognosis of B-CLL. An overview of the sequences disclosed by the present invention is present here:

SEQ ID No:1 includes 20.000 bp human genome sequence, derived from BAC clone acc. no. AC063949. It includes the region encoding the mRNAs corresponding to the cDNAs described in Seq ID Nos:2, 4, 6, 7, 8, 9, 10, and 11 and possible up- and down-stream regulatory sequences. Seq ID No:1 includes a subset of the sequence described in Seq ID No:5. The sequence is derived from human 12q21-22.

SEQ ID No:2 includes a cDNA corresponding to a putative mRNA transcript that includes the region encoding the peptide sequence in Seq ID No:3 and up- and downstream regions. It is transcribed from the +strand of Seq ID No:1 and Seq ID No:5 from human 12q21-22. From position 2317, this sequence is identical to Seq ID No:4.

Seq ID No:3 contains a peptide sequence encoded by some mRNAs transcribed from the region on human chromosome 12q21-22 included in Seq ID No:1 and Seq ID No:5. It is encoded by the mRNA sequences identified as cDNAs in Seq ID Nos 02, 04, 09 and 11.

Seq ID No:4 includes a cDNA corresponding to a putative mRNA transcript that includes the region encoding the peptide sequence in Seq ID No:3 and up- and downstream regions. It is transcribed from the +strand of Seq ID No:1 and Seq ID No:5 from human 12q21-22.

Seq ID No:5 includes 80.000 bp human genome sequence, derived from BAC clone acc. no. AC063949. It is an expansion of the genomic sequence included in Seq ID No:1. which is contained within this sequence. It includes the region encoding the mRNAs corresponding to the cDNAs described in Seq ID Nos 02, 04, 06, 07, 08, 09, 10 and 11 and possible up- and down-stream regulatory sequences. The sequence is derived from human 12q21-22.

Seq ID No:6 corresponds to a cDNA detected by cDNA cloning, corresponding to an mRNA transcript. It includes two exons. It is transcribed from the +strand of Seq ID No:1 and Seq ID No:5 from human 12q21-22.

Seq ID No:7 corresponds to a cDNA detected by cDNA cloning, corresponding to an mRNA transcript. It includes three exons, the first and third are identical to the two exons in Seq ID No:6. It is transcribed from the +strand of Seq ID No:1 and Seq ID No:5.doc from human 12q21-22.

Seq ID No:8 corresponds to a cDNA detected by cDNA cloning, corresponding to an mRNA transcript. It includes two exons, the last is also present as exon-2 in a human cDNA clone

(sequence acc. no. BC036936) (Seq ID No:9.doc). It is transcribed from the +strand of Seq ID No:5.doc from human 12q21-22.

Seq ID No:9 corresponds to a human cDNA sequence (sequence acc. no. BC036936). It is transcribed from the +strand Seq ID No:5 from 12q21-22. We have not cDNA cloned this cDNA, but a splice variant (Seq ID No:8), where exon-2 of this sequence was spliced to exon-1 of Seq ID No:6 was detected by cDNA cloning. It is transcribed from the +strand of Seq ID No:5 from human 12q21-22.

Seq ID No:10 corresponds to a cDNA detected partly by cDNA cloning, partly by PCR analysis, corresponding to an mRNA transcript. It includes two exons, exon-1 includes the region encoding Seq ID No:3 and exon-1 from Seq ID No:6; exon-2 is identical to exon-2 in Seq ID No:6 and exon-3 in Seq ID No:7. It is transcribed from the +strand of Seq ID No:1.doc and Seq ID No:5 from human 12q21-22.

Seq ID No:11 corresponds to a cDNA detected by cDNA cloning, corresponding to an mRNA transcript. It includes one exon. The sequence includes the the region that encodes Seq ID No:3, exons 2 and 3 from Seq ID No:7 and the region between those exons. It is transcribed from the +strand of Seq ID No:1 and Seq ID No:5 from human 12q21-22.

Seq ID No:12 Is an exon sequence. It corresponds to the first exon in Seq ID No: 9 It is transcribed from the +strand of Seq ID No:5 from human 12q21-22.

Seq ID No:13 Is an exon sequence. It corresponds to the first exon in Seq ID No: 6, 7, 8, and it is included in Seq ID No: 2. It is transcribed from the +strand of Seq ID No:1 and 5 from human 12q21-22.

Seq ID No:14 Is an exon sequence. It is identical to Seq ID No:13, but with an additional GT dinucleotide at the 3'end, caused by the use of an alternative splice site. It can replace Seq ID No:13 as the first exon in Seq ID No: 6, 7, 8, and be included in Seq ID No: 2. It is transcribed from the +strand of Seq ID No:1 and 5 from human 12q21-22.

Seq ID No:15 Is an exon sequence. It corresponds to the second exon in Seq ID No:7 and it is included in Seq ID No: 2, 4, 10 and 11. It is transcribed from the +strand of Seq ID No:1 and 5 from human 12q21-22.

Seq ID No:16 Is an exon sequence. It corresponds to the third exon in Seq ID No:7, it is the second exon in Seq ID No:6 and 11 and it is included in Seq ID No:2 and 4. It is transcribed from the +strand of Seq ID No:5 from human 12q21-22.

Seq ID No:17 Is the sequence encoding the peptide in Seq ID No: 3. It is included in Seq ID No:2, 4, 10 and 11. It is transcribed from the +strand of Seq ID No:5 from human 12q21-22.

Seq ID No:18 Is an exon sequence. It corresponds to the second exon in Seq ID No:8 and 9. It is transcribed from the +strand of Seq ID No:5 from human 12q21-22.

Methods of diagnosis

One important aspect of the present invention relates to diagnosis of a subtype of B-cell chronic lymphocytic leukaemia (B-CLL) having poor prognosis. A further important aspect of the invention relates to prognosis of B-CLL. These methods are based on the discovery by the present inventors that an expression product which comprise at least one nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 is (are) only present in particular subtypes of B-CLL associated with poor prognosis and completely absent in other subtypes of B-CLL and in healthy tissue (see in particular example 2). By completely absent is meant that the expression products are not detected in any of the other tissue types with the methods used in the appended examples. This is indicative of a complete absence of any transcript or a very low level of transcript in the other tissue types.

The expression product is encoded by SEQ ID No 1 and/or SEQ ID No 5, and the expression product is selected from the group consisting of transcriptional products and translational products.

Thus, the present invention relates to a method for detecting the presence or absence of at least one expression product, wherein the at least one expression product comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in a biological sample isolated from an individual for establishing a differential diagnosis of B-CLL or for determining the prognosis of the B-CLL.

"Expression product" is herein meant to be a product which is the result of the expression of a polynucleotide such as DNA sequence, e.g. a genomic DNA sequence, and is in the form of either a polypeptid or in the form of a polynucleotide, i.e. an expression product can be selected from the group consisting of a transcriptional product and a translational product. In the case where the expression product is a polynucleotide, said polynucleotide is preferably mRNA selected from the group consisting of mRNA, pre-mRNA, pre-pro-mRNA .

A "transcriptional product" or a "transcription product" is herein meant to be a product resulting from a transcription of a polynucleotide such as a DNA molecule, preferably a genomic DNA molecule. A transcriptional product is inherently a nucleotide, such as an oligonucleotide or a polynucleotide.

A "translational product" or a "translation product" is herein meant to be a product resulting from a translation of a transcriptional product such as a mRNA. A translational product is inherently a oligopeptide or a polypeptide.

The expression product of the present invention has almost exclusively been found as transcription products in patients with poor B-CLL prognosis. Based on the experimental data presented in the herein, the inventors expect that it turns out that the subtype of B-CLL is characterised solely or better by the presence of a transcriptional or translational product which comprise a sequence selected from the group consisting of SEQ ID No:3, SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18.

Preferably the individual is a mammal, more preferably a human. It is also expected that the gene encoded by SEQ ID No 1 and/or SEQ ID No 5 and the expression products derived from said gene can be used as a diagnostic tool in other species in particular in mammals selected from the group: domestic animals such as cow, horse, sheep, pig; and pets such as cat or dog.

In the case that the expression product is a transcriptional product, this transcriptional product just needs to comprise at least one of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18. These transcriptional products will preferably be derived from SEQ ID No 1 and/or SEQ ID No 5 and may be in the form of mRNA or any pre- or pro-forms of said mRNA. As described, the transcriptional product may comprise at least one of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18, such as one of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18, such as two of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18, such as 3 of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18, e.g. 4 of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18, such as 5 of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 and ultimately the transcriptional product may comprise all of the six nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18. The transcriptional product of the present invention can have any sequence which is a result of combining the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 as long as the specific nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15,

SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 can be identified as intact sequences in the transcriptional product.

Examples of transcriptional products in the form of specific mRNAs which comprise at least one of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 are the sequences corresponding to SEQ ID No 2 (short cDNA clone) SEQ ID No 4 (long cDNA clone) SEQ ID No 6, SEQ ID No 7, SEQ ID No 8, SEQ ID No 9, SEQ ID No 10 and SEQ ID No 11. These mRNA sequences have been found in patients with poor prognosis.

It is obvious for a person skilled in the art that any fragments of SEQ ID No:2, SEQ ID No:4, SEQ ID No:6, SEQ ID No:7, SEQ ID No:8, SEQ ID No:9, SEQ ID No:10 and SEQ ID No:11 will have the same diagnostic value as long as the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 can be found in the fragments.

The mRNA sequence may be detected in a sample using hybridisation techniques. In particular when more than one analysis is to be performed at the same time it is advantageous to use a DNA array comprising e.g. an oligomer of at least 15 consecutive bases selected from the group consisting of SEQ ID No:2, SEQ ID No:4, SEQ ID No:6, SEQ ID No:7, SEQ ID No:8, SEQ ID No:9, SEQ ID No:10, SEQ ID No:11, SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18.

Another way of detecting the presence or absence of the transcriptional product is by specifically amplifying the transcriptionals product having a sequence corresponding to SEQ ID No 2, 4, 6, 7, 8, 9, 10 or 11 or fragments thereof. This can be done by selecting primer pairs which cause only the amplification of these sequences.

Generally, hybridisation techniques are selected from not limited to the group consisting of in situ hybridisation, northern blots, Southern blots, dot blots and PCR based techniques.

A non-limiting list of PCR based techniques include rt-PCR, quantitative PCR and realtime PCR.

According to yet another embodiment, the translational product is a protein encoded by a polynucleotide selected from the group consisting of SEQ ID No:1, SEQ ID No:5, SEQ ID No:2, SEQ ID No:4, SEQ ID No:6, SEQ ID No:7, SEQ ID No:8, SEQ ID No:9, SEQ ID No:10 and SEQ ID No:11. Detection of this protein can be done with state of the art methods including the detection with an antibody directed against said protein, such as Western blotting, more preferably by using a fluorescently labelled antibody, preferably wherein the method comprises the use of flowcytometry, such as FACS. Other methods include but are not limited to gel electrophoresis, gel filtration, ion exchange chromatography, FPLC, mass spectrometry and immunohistochemistry.

Preferably, said protein is selected from the group comprising SEQ ID No 3 (protein), or a protein sharing at least 60 % sequence identity with SEQ ID No 3. The protein with the amino acid sequence set forth in SEQ ID No 3 is the longest open reading frame in the cDNA sequence of SEQ ID No 2 or 4.

In a specific embodiment of the present invention is a method for determining whether an individual has a B-CLL sub-type with poor prognosis, the method comprising determining the level of an expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 of said individual, and indicating the individual as having a B-CLL sub-type with poor prognosis if the level of the expression product is at or beyond a discriminating value and indicating the individual as not having a B-CLL sub-type with poor prognosis if the level of the expression product is not at or beyond the discriminating value, the discriminating value being a value which has been determined by measuring the level of the expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in both a healthy control population and a population with known B-CLL sub-type with poor prognosis, thereby determining said discriminating value which identifies the B-CLL sub-type population having a poor prognosis.

In this method the individual may be a member of an unselected population or be a member of a population already identified as having a B-CLL sub-type with a poor prognosis.

The above method may be performed such that the determination is performed at several time points at intervals as part of a monitoring of a cancer patient after or during the treatment for primary cancer.

The methods described so-far relate to the determination of the presence or absence of an expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18. By quantitatively measuring the amount of an expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in a biological sample isolated from an individual, it is possible to predict the progression of B-CLL in an individual.

In one embodiment the quantitative measurement is performed during treatment to estimate the efficiency of such treatment.

A preferred embodiment of the above diagnostic and prognostic methods is a method for detecting the presence or absence an expression product, wherein said at least one expression product comprise the nucleotide sequence of SEQ ID No:15 in a biological

sample isolated from an individual for establishing a differential diagnosis of B-CLL or for determining the prognosis of the B-CLL.

A further preferred embodiment of the above diagnostic and prognostic methods is a method for detecting the presence or absence an expression product, wherein said at least one expression product comprise the nucleotide sequence of SEQ ID No:16 in a biological sample isolated from an individual for establishing a differential diagnosis of B-CLL or for determining the prognosis of the B-CLL.

Yet a further preferred embodiment of the above diagnostic and prognostic methods is a method for detecting the presence or absence an expression product, wherein said at least one expression product comprise the nucleotide sequence spanning the junction sequence between Exon-2 (SEQ ID No:15) and Exon-3 (SEQ ID No:16) in a biological sample isolated from an individual for establishing a differential diagnosis of B-CLL or for determining the prognosis of the B-CLL.

The nucleotide sequence spanning the junction between Exon-2 and Exon-3 is the last 20 nucleotides of the 3'-end of SEQ ID No:15 and the first 20 nucleotides of the 5'-end of SEQ ID No:16.

The "junction sequence" between two nucleotide sequences, such as two exons, is herein defined as the at least 20 3'-nucleotides of the first exon which is located 5' relative to the second exon and the at least 20 5'-nucleotides of the second exon which is located 3' relative to the first exon.

For all diagnostic applications of the present invention, the biological sample may be selected from the group comprising blood, serum, plasma, urine, saliva, lymph node biopsy, bone marrow, spinal liquid, spleen biopsy, and liver biopsy. The cells to be assessed in a sample are preferably leukocytes, mononuclear leukocytes or lymphocytes or B-lymphocytes.

A further embodiment of the present invention also includes a diagnostic kit for *ex vivo* or *in situ* diagnosis of a subtype of B-cell chronic lymphocytic leukaemia (B-CLL) in a individual, the kit comprising a detector molecule capable of detecting the presence or absence of at least one expression product, wherein said at least one expression product comprise a nucleotide sequence selected from the group consisting of SEQ ID SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in a biological sample isolated from the individual.

The detector molecule is preferably a nucleotide and even more preferably a nucleotide capable of hybridising to a nucleotide sequence selected from the group consisting of SEQ ID SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 under stringent condition.

B-CLL therapy

With the identification of a new sub-type of B-CLL having a poor prognosis, the present inventors also provide methods for treatment of B-CLL in such patients. This method is based on the finding that transcription products comprising these sequence products are present in B-CLL cells of patients with the poor prognosis. By modifying the activity and/or level of these transcription products, a treatment and/or cure for B-CLL is provided.

Accordingly, in a therapeutic aspect of the present invention there is provided a method of treating a B-CLL sub-type with poor prognosis comprising administering to an individual with a poor prognosis B-CLL diagnosis a compound capable of decreasing or inhibiting the formation of an expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18.

One such method is based on administering an oligonucleotide capable of inhibiting transcription from SEQ ID No 1 and/or SEQ ID No 5. Said oligonucleotide may comprises at least 8-10 consecutive nucleotides from the sequence of SEQ ID No 1. These sequences constitute the putative promoter sequences controlling the transcription transcription products which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18. The oligonucleotides bind specifically to the promoter sequences and inhibit transcription of the gene. Such oligonucleotides may comprises nucleotide monomers selected from the group: DNA, RNA, LNA, PNA, methylated DNA, methylated RNA, more preferably PNA or LNA.

In a more preferred embodiment the therapeutic methods comprise administering an oligonucleotide capable of binding to a transcriptional product and preventing translation by destroying the transcriptional product. One particularly preferred embodiment of this aspect is RNA interference (RNAi) oligonucleotides.

The discovery of the phenomenon RNAi has revealed an entirely new level of gene regulation in eukaryotic cells. It is based on the observation that the presence of long double stranded RNA (dsRNA) in a cell almost completely eliminates the expression of the gene having the same sequence, whereas expression of other unrelated genes are left undisturbed. Although this observation had been know for time in plants as posttranscriptional gene silencing (PTGS) it was not before it was characterised as a general mechanism throughout the animal kingdom that its potentials were fully appreciated. Over the last few years it has been developed as a robust technique to knock down any desirable gene in worms and flies, and quickly a large body of information was gathered about the function of genes in these organisms. Due to the activation of the interferon system by long dsRNA the RNAi method was at that time not applicable in a mammalian system.

A key observation that allowed the harnessing of RNAi as a tool for regulating gene expression in mammals was the observation that chemically synthesised oligo-mer small interfering RNAs (siRNA) effectively suppress gene expression in several human cell lines without inflecting interferon response. This has triggered new promises for siRNA as a therapeutic drug in humans.

RNAi works by hybridising specifically to the mRNA transcribed by the cell to form a (partly) double stranded RNA molecule. This is recognised as a double stranded molecule by the cell's own nucleases, which degrade them.

In order for the technique to work efficiently, the siRNA oligonucleotide comprises a sequence of 5-30 consecutive nucleotides which is the complementary sequence of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18. By targeting at least one of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18, the transcriptional products characterised by the nucleotide sequences from the group of SEQ ID No:2, SEQ ID No:4, SEQ ID No:6, SEQ ID No:7, SEQ ID No:8, SEQ ID No:9, SEQ ID No:10 and SEQ ID No:11 will be eliminated. Example 5 shows that cells characteristic for the poor prognosis B-CLL sub-type can be eliminated by destroying the herein mentioned transcription products.

By "complementarity" is meant a nucleic acid that can form hydrogen bond(s) with other DNA or RNA sequence by either traditional Watson-Crick or other non-traditional types of base-paired interactions, e.g. Hoogsteen type.

Preferred siRNA molecules of the present invention are between 5 to 30 nucleotides long, such as 8-30 nucleotides long, such as 8-25 nucleotides, e.g. 8-24 nucleotides, e.g. 8-23 nucleotides, e.g. 8-22 nucleotides, e.g. 8-21 nucleotides, such as 8-20 nucleotides, e.g. 9-23 nucleotides, e.g. 10-23 nucleotides, such as 11-23 nucleotides, e.g. 12-23 nucleotides such as 13-23 nucleotides, e.g. 14-23 nucleotides, e.g. 15-23 nucleotides, such as 16-23 nucleotides, such as 8 nucleotides, 9 nucleotides, 10 nucleotides, 11 nucleotides, 12 nucleotides, 13 nucleotides, 14 nucleotides, 15 nucleotides, 16 nucleotides, 17 nucleotides, 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, 25 nucleotides, 26 nucleotides, 27 nucleotides, 28 nucleotides, 29 nucleotides, such as 30 nucleotides long.

RNAi oligonucleotides may be administered to the cell, or a vector may be transfected into the cells, said vector comprising a promoter region capable of directing the expression of at least one RNAi oligonucleotide. Due to the very restricted expression of the AMB-1 gene, it is not important only to target the RNAi oligos or the vectors to B-CLL cells.

One way of targeting to blood cells comprises using a heparin receptor for targeting to blood cells.

Another way of targeting the transcriptional products which comprise at least one nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 is to use an antisense construct comprising a promoter sequence capable of directing the transcription of at least part of the antisense equivalent of SEQ ID No 1 or 2 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 specifically to the poor prognosis B-CLL sub-type.

When desired targeting to B-CLL cells can be performed using the CD19 or CD20 receptor. The CD19 receptor is particularly preferred since it internalises its ligand.

In a further therapeutic embodiment the compound is a gene therapy vector comprising a promoter sequence operably linked to a sequence coding for a protein capable of inhibiting cell division in the cell and/or capable of killing the cell, said promoter sequence being a tissue specific promoter capable of directing expression only in B cells, more preferably only in B-CLL cells. One particularly preferred promoter sequence is the extremely cell specific promoter of SEQ ID No:1 or SEQ ID No:5. When this promoter is used targeting of the suicide vector is not very important, since it will only be active in the cells in which AMB-1 is expressed and these are the cells to be targeted by the suicide gene.

Deletion studies will determine the exact length of the promoter sequence counted from the transcription start site. Accordingly, the promoter may comprise at least 100 nucleotides of Seq_ID:1 or Seq_ID:5, such as at least 200 nucleotides, for example at least 300 nucleotides, such as at least 400 nucleotides, for example at least 500 nucleotides, such as at least 600 nucleotides, for example at least 700 nucleotides, such as at least 800 nucleotides, for example at least 900 nucleotides, such as at least 1000 nucleotides, for example at least 1100 nucleotides, such as at least 1200 nucleotides, for example at least 1300 nucleotides, such as at least 1400 nucleotides, for example at least 1500 nucleotides, such as at least 1600 nucleotides, for example at least 1700 nucleotides, such as at least 1800 nucleotides, for example at least 1900 nucleotides, such as at least 2000 nucleotides, for example at least 2500 nucleotides, such as at least 3000 nucleotides, for example at least 3500 nucleotides, such as at least 5000 nucleotides, for example at least 10,000 nucleotides.

The specificity of expression of mRNAs described by the present invention are striking. The RT-PCR data and the Northern blot data using the dot-blot disclose that the mRNAs of the present invention are expressed either at very low levels in other tissues or only in the B-CLL patients where one can detect it.

Thus one embodiment of the present invention is the use of the promotor region for use in gene therapy. The promotor is defined as any sequence within SEQ ID No:1 and SEQ ID No:5 that directs the formation of an expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18, said expression

product being any of the sequences selected from the group consisting of SEQ ID No:2, SEQ ID No:4, SEQ ID No:6, SEQ ID No:7, SEQ ID No:8, SEQ ID No:9, SEQ ID No:10 and SEQ ID No:11 in B-CLL cells or any other cell or tissue types in which any of the sequences are transcribed.

One way that one can use the promotor region in gene therapy is to make a gene therapy construct where the promotor drives the expression of a cell suicide gene such as but not limited to the gene for HSV-1 thymidine kinase, the varicella-zoster, virus thymidine kinase gene, E.Coli cytosine deaminase, the nitroreductase gene or the E.Coli Deo gene (Yazawa et al. & Kirn D. et al.). This would allow for selective expression of the suicide genes in B-CLL cells.

Alternatively, the promotor could be used for a selective expression of genes that could have curative effects when expressed in B-CLL cells, but unwanted effects if expressed ubiquitously.

Also, one embodiment of the present invention relates to the use of the promotor region for use in screening assays where the promotor is linked to a reporter gene and transfected into B-CLL cells in which the reporter gene will be expressed. This approach would allow for easy screening for compounds that would turn off the expression of the reporter gene for example by killing the cell.

A presently preferred embodiment relates a gene therapy vector of the present invention comprising an oligonucleotide capable of inhibiting transcription from SEQ ID No 1 and/or SEQ ID No 5, wherein the promoter is a B-CLL specific promoter, which may or may not be operably linked to a protein selected from the group comprising HSV-1 thymidine kinase, the varicella-zoster, virus thymidine kinase gene, E.Coli cytosine deaminase, the nitroreductase gene or the E.Coli Deo gene.

In one embodiment the compound is a therapeutic antibody directed against a polypeptide having the amino acid sequence of SEQ ID No 3, preferably wherein said antibody is a human or humanised antibody. Another possibility is to identify a modulator of binding of SEQ ID No 3 to its receptor within or outside the cell and to administer this modulator to the cells.

4-helical cytokines

A further object of preferred embodiments of the present invention is an isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

i) an amino acid sequence of SEQ ID NO: 3,

ii) an amino acid sequence having at least 60% sequence identity compared to the full length sequence of SEQ ID NO:3

ii) a fragment of SEQ ID NO:3 having at least 60% sequence identity compared to the full length sequence of SEQ ID NO:3.

The protein encoded by SEQ ID No 1 and/or SEQ ID No 5 shares a very small sequence identity with any known protein. However, it has been possible to use 2D and 3D analytical tools to identify the protein as a 4-helical cytokine. The 3D structure of the protein is very similar to 4-helical cytokines and in particular to IL4.

IL4 is a very important cytokine in B-CLL biology. IL4 is not expressed by B-CLL cells, but the IL4 receptor is found on the cells. The IL4 that stimulates B-CLL cells is believed to be produced by T-lymphocytes. The role of IL4 in B-CLL biology is complicated. It has been suggested that IL4 can inhibit B-CLL DNA synthesis and proliferation. Other reports demonstrated that IL4 protects B-CLL cells from apoptosis by upregulating Bcl-2, and IL4 was shown to inhibit apoptosis without stimulating proliferation. Recently, a clinical study in Sweden has confirmed these in vitro studies since IL4 administration to B-CLL patients resulted in increased numbers of B-CLL cells in the blood, suggesting that IL4 had a stimulatory or anti-apoptotic effect on the B-CLL cells in vivo (Lundin, et al.).

In many systems the effects of IL13 are largely similar to those of IL4, but IL13 is slightly less potent that IL4. It is unclear whether B-CLL cells express IL13, but the cells do express the IL13 receptor. The effects of IL13 in B-CLL are controversial. While Chaouchi et al. suggested that IL13, like IL4 protects B-CLL cells from apoptosis (Chaouchi et al), studies by Fluckiger et al. suggest that this is not the case (Fluckiger et al.).

The combined finding of 2D and 3D structure similarity to 4-helical cytokines and the importance of IL4 in B-CLL strongly suggests that the novel class of proteins of which the AMB-1 protein is one representative are cytokines.

These polypeptides constitute a novel class of proteins sharing 2D and 3D structure similarities with 4-helical cytokines. In a preferred embodiment, the isolated polypeptide comprises or essentially consists of the amino acid sequence of SEQ ID No. 3 or a fragment thereof. This particular protein at least can be used for diagnosis, for raising antibodies for use in therapy against B-CLL, and for protective or therapeutic immunisation of an individual against B-CLL.

Consequently, the isolated polypeptide preferably has interleukin activity or interleukin like activity, such as having IL3, IL13, GM-CSF, TGF-β, IGF activity, more preferably having IL4 activity or IL4 like activity.

Probably the isolated polypeptides are capable of forming homo- or hetero-oligomer with each other and among themselves. Such oligomers are also within the scope of the present invention. Such oligomers may comprise at least one isolated polypeptides as defined in any the present invention, such as a dimer, a trimer, a quatramer, a quintamer, a

hexamer, an octamer, a decamer, a dodecamer. In biological systems the activity may be attributed only to dimer or higher –mer.

The protein defined by SEQ ID No 3 shares very little sequence identity with known cytokines and interleukines and as a matter of fact very little sequence identity with any known protein. Consequently the present inventors contemplates that the group comprises functionally equivalent polypeptide sharing at least 60% sequence identity with SEQ ID No 3, more preferably at least 70% sequence identity, more preferably at least 80 % sequence identity, such as at least 90 % sequence identity, for example at least 95 % sequence identity, such as at least 97 % sequence identity, for example at least 98 % sequence identity.

Activity as a cytokine or interleukin can be assessed in a biological assay where the polypeptide is contacted with a cytokine dependent cell line. Accordingly, polypeptides with cytokine or interleukin like activity can also be identified by similar methods.

One approach to assess cytokine/interleukin activity in a biological assay is to express the CDS (SEQ ID No:17) reading frame in a baculovirus system (Invitrogen, Carlsbad, USA) and purify the protein . The recombinant protein can be assayed in cytokine induced proliferation assays as described in general in the eBioscience catalog & Reference Manual 2002 p.260-262 (eBioscience, San Diego, USA). In particular IL4 activity can be determined using the CTh4S cell line as described by Petersen et al (see Example 8).

The promoter sequence (which forms part of SEQ ID No 1 and/or SEQ ID No 5) and the coding sequences (SEQ ID No:3) can be used in various aspects of gene therapy and immunotherapy.

Further polynucleotide sequences from other individuals or other species with the same function can be isolated by one of the following methods, which each form independent aspects of the present invention.

A first method for identifying a nucleotide sequence encoding a 4-helical cytokine comprises the steps of:
i) isolating mRNA from a biological sample,
ii) hybridising the mRNA to a probe comprising at least 10 nucleotides of the coding sequence of SEQ ID No 1 and/or SEQ ID No 5 under stringent conditions,
iii) determining the nucleotide sequence of a sequence capable of hybridising under step ii), and
iv) determining the presence of an open reading frame in the nucleotide sequence determined under step iii).

A second method for identifying a nucleotide sequence encoding a 4-helical cytokine is a computer assisted method comprising the steps of

i) performing a sequence similarity search of at least 10 nucleotides of the coding sequence SEQ ID No 1 and/or SEQ ID No 5,
ii) aligning "hits" to said coding sequence,
iii) determining the presence of an open reading frame in the "hits".

It is highly likely that other similar polypeptides encoding further 4-helical cytokines can be found in other individuals and/or other species of mammals. In particular, individuals of other geographical origin may carry genes which differ from the polynucleotides of the present invention. It is also conceivable that similar sequences can be found in closely and even in distantly related species.

Functional Equivalents

Modification and changes may be made in the structure of the peptides of the present invention and DNA segments which encode them and still obtain a functional molecule that encodes a protein or peptide with desirable characteristics. The following is a discussion based upon changing the amino acids of a protein to create an equivalent, or even an improved, second-generation molecule. The amino acid changes may be achieved by changing the codons of the DNA sequence, according to the genetic code.

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies, binding sites of receptors, or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporate herein by reference). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been assigned a hydropathic index on the basis of their hydrophobicity and charge characteristics (Kyte and Doolittle, 1982), these are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, ie. still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ±2 is preferred, those which are within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U.S. Pat. No. 4,554,101, incorporated herein by reference, states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U.S. Pat. No. 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0±1); glutamate (+3.0±1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5±1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ±2 is preferred, those which are within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions which take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

Functional equivalents and variants are used interchangably herein. In one preferred embodiment of the invention there is also provided variants of a 4-helical cytokine, and variants of fragments thereof. When being polypeptides, variants are determined on the basis of their degree of identity or their homology with a predetermined amino acid sequence, said predetermined amino acid sequence being SEQ ID No. 3 or a fragment thereof.

Accordingly, variants preferably have at least 60 % sequence identity, for example at least 65% sequence identity, such as at least 70 % sequence identity, for example at least 75% sequence identity, for example at least 80% sequence identity, such as at least 85 % sequence identity, for example at least 90 % sequence identity, such as at least 91 % sequence identity, for example at least 91% sequence identity, such as at least 92 % sequence identity, for example at least 93 % sequence identity, such as at least 94 % sequence identity, for example at least 95 % sequence identity, such as at least 96 %

sequence identity, for example at least 97% sequence identity, such as at least 98 % sequence identity, for example 99% sequence identity with the predetermined sequence.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity # of identical positions/total # of positions (e.g., overlapping positions) x 100). In one embodiment the two sequences are the same length. The sequence identity is preferably calculated relative to the full length sequence of the molecule of the present invention.

A degree of "sequence identity" of amino acid sequences is a function of the number of identical amino acids at positions shared by the amino acid sequences. A degree of "sequence homology" or "sequence similarity" of amino acid sequences is a function of the number of amino acids, i.e. structurally related, at positions shared by the amino acid sequences. Sequence identity is determined by the alignment algorithm that performs global alignments which has been described by Smith TF and Waterman MS (Smith TF et al.)

A list of the standard qualifiers and the default values for the alignment algorithm is given below:

Standard (Mandatory) qualifiers		**Allowed values**	**Default**
[-asequence] (Parameter 1)	Sequence USA	Readable sequence	**Required**
[-bsequence] (Parameter 2)	Sequence database USA	Readable sequence(s)	**Required**
-gapopen	The gap open penalty is the score taken away when a gap is created. The best value depends on the choice of comparison matrix. The default value assumes you are using the EBLOSUM62 matrix for protein sequences, and the EDNAFULL matrix for nucleotide sequences.	Number from 1.000 to 100.000	10.0 for any sequence

-gapextend	The gap extension penalty is added to the standard gap penalty for each base or residue in the gap. This is how long gaps are penalized. Usually you will expect a few long gaps rather than many short gaps, so the gap extension penalty should be lower than the gap penalty. An exception is where one or both sequences are single reads with possible sequencing errors in which case you would expect many single base gaps. You can get this result by setting the gap open penalty to zero (or very low) and using the gap extension penalty to control gap scoring.	Number from 0.100 to 10.000	0.5 for any sequence
[-outfile] (Parameter 3)	Output alignment file name	Alignment output file	
Additional (Optional) qualifiers		**Allowed values**	**Default**
-datafile	This is the scoring matrix file used when comparing sequences. By default it is the file 'EBLOSUM62' (for proteins) or the file 'EDNAFULL' (for nucleic sequences). These files are found in the 'data' directory of the EMBOSS installation.	Comparison matrix file in EMBOSS data path	EBLOSUM62 for protein EDNAFULL for DNA
Advanced (Unprompted) qualifiers		**Allowed values**	**Default**
-[no]brief	Brief identity and similarity	Boolean value Yes/No	Yes

An "unrelated" or "non-homologous" sequence shares less than 40% identity, though preferably less than 25% identity, with one of the 4-helical cytokine sequences of the present invention. The term "substantial identity" means that two peptide sequences, when optimally aligned, such as by the Smith and Waterman algorithm using default gap weights, share at least 80 percent sequence identity, preferably at least 90 percent sequence identity, more preferably at least 95 percent sequence identity or more (e.g., 99 percent sequence identity). Preferably, residue positions which are not identical differ by conservative amino acid substitutions.

Additionally, variants are also determined based on a predetermined number of conservative amino acid substitutions as defined herein below. Conservative amino acid substitution as used herein relates to the substitution of one amino acid (within a predetermined group of amino acids) for another amino acid (within the same group), wherein the amino acids exhibit similar or substantially similar characteristics.

Within the meaning of the term "conservative amino acid substitution" as applied herein, one amino acid may be substituted for another within the groups of amino acids indicated herein below:

- Amino acids having polar side chains (Asp, Glu, Lys, Arg, His, Asn, Gln, Ser, Thr, Tyr, and Cys,)
- Amino acids having non-polar side chains (Gly, Ala, Val, Leu, Ile, Phe, Trp, Pro, and Met)
- Amino acids having aliphatic side chains (Gly, Ala Val, Leu, Ile)
- Amino acids having cyclic side chains (Phe, Tyr, Trp, His, Pro)
- Amino acids having aromatic side chains (Phe, Tyr, Trp)
- Amino acids having acidic side chains (Asp, Glu)
- Amino acids having basic side chains (Lys, Arg, His)
- Amino acids having amide side chains (Asn, Gln)
- Amino acids having hydroxy side chains (Ser, Thr)
- Amino acids having sulphor-containing side chains (Cys, Met),
- Neutral, weakly hydrophobic amino acids (Pro, Ala, Gly, Ser, Thr)
- Hydrophilic, acidic amino acids (Gln, Asn, Glu, Asp), and
- Hydrophobic amino acids (Leu, Ile, Val)

Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.

Accordingly, a variant or a fragment thereof according to the invention may comprise, within the same variant of the sequence or fragments thereof, or among different variants of the sequence or fragments thereof, at least one substitution, such as a plurality of substitutions introduced independently of one another.

It is clear from the above outline that the same variant or fragment thereof may comprise more than one conservative amino acid substitution from more than one group of conservative amino acids as defined herein above.

The addition or deletion of at least one amino acid may be an addition or deletion of from preferably 2 to 250 amino acids, such as from 10 to 20 amino acids, for example from 20 to 30 amino acids, such as from 40 to 50 amino acids. However, additions or deletions of more than 50 amino acids, such as additions from 50 to 100 amino acids, addition of 100 to 150 amino acids, addition of 150-250 amino acids, are also comprised within the present invention. The deletion and/or the addition may - independently of one another - be a deletion and/or an addition within a sequence and/or at the end of a sequence.

The polypeptide fragments according to the present invention, including any functional equivalents thereof, may in one embodiment comprise less than 250 amino acid residues, such as less than 240 amino acid residues, for example less than 225 amino acid residues, such as less than 200 amino acid residues, for example less than 180 amino acid residues, such as less than 160 amino acid residues, for example less than 150 amino acid residues, such as less than 140 amino acid residues, for example less than 130 amino acid residues, such as less than 120 amino acid residues, for example less than 110 amino acid residues, such as less than 100 amino acid residues, for example less than 90 amino acid residues,

such as less than 85 amino acid residues, for example less than 80 amino acid residues, such as less than 75 amino acid residues, for example less than 70 amino acid residues, such as less than 65 amino acid residues, for example less than 60 amino acid residues, such as less than 55 amino acid residues, for example less than 50 amino acid residues.

"Functional equivalency" as used in the present invention is according to one preferred embodiment established by means of reference to the corresponding functionality of a predetermined fragment of the sequence.

Functional equivalents or variants of a 4-helical cytokine will be understood to exhibit amino acid sequences gradually differing from the preferred predetermined 4-helical cytokine, as the number and scope of insertions, deletions and substitutions including conservative substitutions increases. This difference is measured as a reduction in homology between the preferred predetermined sequence and the fragment or functional equivalent.

All fragments or functional equivalents of SEQ ID No. 3 are included within the scope of this invention, regardless of the degree of homology that they show to the respective, predetermined 4-helical cytokines disclosed herein. The reason for this is that some regions of the 4-helical cytokines are most likely readily mutatable, or capable of being completely deleted, without any significant effect on the binding activity of the resulting fragment.

A functional variant obtained by substitution may well exhibit some form or degree of native cytokine activity, and yet be less homologous, if residues containing functionally similar amino acid side chains are substituted. Functionally similar in this respect refers to dominant characteristics of the side chains such as hydrophobic, basic, neutral or acidic, or the presence or absence of steric bulk. Accordingly, in one embodiment of the invention, the degree of identity is not a principal measure of a fragment being a variant or functional equivalent of a preferred predetermined fragment according to the present invention.

One particularly preferred method of determining the degree of functional equivalence is by performing a biological or chemical assay such as the assays described in the appended examples. Preferred functional equivalents of SEQ ID No 3 are those that have a K_D with respect to a predefined receptor which is less than 10 times higher than the K_D of the polypeptide of SEQ ID No 1 with respect to the same receptor, more preferably less than 5 times higher, more preferably less than 2 times higher.

With respect to functional equivalence this may be defined in a biological assay based on a cytokine dependent or stimulated cell line. Such cell lines are e.g. available from American Type Culture Collection, P.O.Box 1549, Manassas, VA 20108 USA. The following cell lines at least are available for testing cytokines and in particular interleukins:

Accession number	Description	Activity
CRL-1841	**TH-2 clone A5E**	**IL2 dependent, IL4 stimulated**
CRL-2003	**TF-1**	**IL3 dependent**
CRL-2407	**NK-92**	**IL2 dependent**
CRL-2408	**NK-92MI**	**IL2 dependent**
CRL-2409	**NK92CI**	**IL2 dependent**
CRL-9589	**AML-193**	**IL3 stimulated, GM-CSF dep.**
CRL-9591	**MV-4-11**	**GM-CSF dependent**
TIB-214	**CTLL-2**	**IL2 dependent**

The following cell lines are available from DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, D-38124 Braunschweig, GERMANY. As can be seen from the table, some of the cell lines can be used to broadly assess cytokine activity whereas others are only reported to respond to one or a few specific cytokines.

Accession number	**Description**	**Acvitity**
ACC 211	Mouse hybridoma, B9	IL6 dependent
ACC 137	Human acute myeloid leukemia, UT-7	Constitutively cytokine responsive to various cytokines.
ACC 104	Human acute megakaryoblastic leukemia	Respond with proliferation to: GM-CSF, IFN-alpha, IFN-á, IFN-gamma, IL2, IL3, IL4, IL6, IL15, NG F, SCF, TNF-alpha, TPO
ACC 247	Human acute myeloid leukemia, OCI-AML5	G-GSF, GM-CSF, IL3, FTL3-ligand
ACC 271	Human acute myeloid leukemia, MUTZ-2	IL3, SCF, G-CSF, M-CSF, IFN-gamma
ACC 334	Human erythroleukemia, TF-1	GM-CSF, IFN-gamma, IL3, IL4, IL5, IL6, IL13, LIF, NGF, OSM, SCF, TNF-alpha, and TPO

The TF-1 cell line mentioned above can be used for assaying IL13 function. This cell line is sensitive to various different cytokines but gives a very strong proliferative response when exposed to IL13. This cell line can in particular be used if there is no response in the IL4 sensitive cell line (CT.h4S). Further cell lines which can be used for distinguishing between IL4 and IL13 activity include cell lines/hybridomas such as B-9-1-3 (Bouteiller et al.) and A201.1 (Andrews et al.).

Pharmaceutical uses of isolated polypeptides

Apart from being used for diagnosis, it is also within the scope of the present invention to use an isolated polypeptide as defined in the invention for a pharmaceutical composition together with a pharmaceutically acceptable carrier. Such pharmaceutical compositions may be used for any of the purposes for which cytokines and in particular interleukin is used at present.

Examples of such uses include the treatment of bone disorders, inflammation, for lowering blood serum cholesterol, allergy, infection, viral infections, hematopoietic disorders, preneoplastic lesions, immune related diseases, autoimmune related diseases, infectious diseases, tuberculosis, cancer, viral diseases, septic shock, reconstitution of the haematopoietic system, induction of the granulocyte system, pain, cardial dysfunction, CNS disorders, depression, artheritis, psoriasis, dermatitis, collitis, Crohn's disease, diabetes, in an individual in need thereof.

It is also within the scope of the present invention to use an isolated polypeptide according to the invention as an adjuvant or as an immune anhancer, for regulating TH2 immune responses, and for suppressing Th1 immune responses.

A further use of an isolated polypeptide of the invention is as a growth factor for administration to cell cultures or as a growth factor for veterinary use, e.g. for stimulating the growth of livestock.

Immunotherapy

Having identified a transcriptional and/or translational product of SEQ ID No 1 and/or SEQ ID No 5 as an etiological factor in B-CLL it is also within the scope of the present invention to perform an immunisation of a patient in need thereof against B-CLL, wherein the immunisation generates an immune response in the patient which recognises a translational product of SEQ ID No 2, SEQ ID No 4, SEQ ID No 6, SEQ ID No 7, SEQ ID No 8, SEQ ID No 9, SEQ ID No 10 and SEQ ID No 11. A preferred immunotherapy is a vaccination against B-CLL by immunising an individual against a translational product of SEQ ID No 1 and/or SEQ ID No 5. In this way the individual builds up antibodies directed against said translational product and any developing B-CLL will be stopped by these antibodies.

Immunisation may be performed in various ways, such as by immunising said individual with at least one isolated polypeptide as defined the present invention and optionally adjuvants and carriers or immunising with an expression construct capable of expressing an isolated polypeptide according to the invention in the cells (DNA vaccination).

Another method comprises peptide loading of dendritic cells, or ex vivo expansion and activation of T-cells, or inducing a CTL response that targets cells expressing the polypeptide encoded by SEQ ID No 1 and/or SEQ ID No 5.

Antibodies

Antibodies against any of the polypeptides belonging to the novel class of proteins identified by the present inventors can be produced by any known method of immunisation.

In one embodiment, the antibodies are produced in a non-human mammal, or in an insect. If antibodies are to be used for therapy in human beings they are preferably subsequently humanised. In one embodiment, the antibody is formulated into a single-chain antibody.

In another embodiment, in particular for therapeutic purposes, the host organism is a human being and the antibody is subsequently produced recombinantly in a non-human mammal, such as a mouse. The antibody may also be produced as a monoclonal antibody in a hybridoma. One way of producing a monoclonal antibody is described in US 5,681,729 in which a human lymphocyte producing an antibody is generated by the steps, in the order mentioned, comprising

1. transplanting human lymphocytes to a mouse lacking both functional T and B cells so that said human lymphocytes take in said mouse's body;
2. immunizing said mouse with a desired antigen so as to generate human lymphocytes producing an antibody specific to said antigen;
3. administering to said mouse an antiserum to mouse cells;
4. recovering lymphocyte containing cells from said mouse;
5. separating human lymphocytes from the recovered cells by centrifugation; and
6. separating said human lymphocytes producing said antibody.
7. immortalizing said human lymphocytes
8. cloning the obtained immortalized human-derived lymphocytes producing said antibody; and

recovering a monoclonal antibody specific to said desired antigen from the cloned immortalized human-derived lymphocytes.

The antibodies of the present invention may be provided as part of a pharmaceutical composition. Such a pharmaceutical composition may be used for treating cancer, preferably for treating leukaemia, more preferably for treating B-CLL leukaemia, more preferably for treating poor prognosis B-CLL leukaemia.

Use of antibodies in therapy

Antibodies directed against epitopes can be used for prevention and/or therapy of for example B-CLL. Antigenic epitopes can be used in vaccines to stimulate an immunological response in a mammal that is directed against cells having the B-CLL-associated epitope found in the AMB-1 protein(s) or functional equivalents. Antibodies directed against the antigenic epitopes of the invention can combat or prevent B-CLL.

An antigenic epitope may be administered to the mammal in an amount sufficient to stimulate an immunological response against the antigenic epitope. The antigenic epitope

may be combined in a therapeutic composition and administered in several doses over a period of time that optimizes the immunological response of the mammal. Such an immunological response can be detected and monitored by observing whether antibodies directed against the epitopes of the invention are present in the bloodstream of the mammal.

Such antibodies can be used alone or coupled to, or combined with, therapeutically useful agents. Antibodies can be administered to mammals suffering from any B-CLL that displays the B-CLL-associated epitope. Such administration can provide both therapeutic treatment, and prophylactic or preventative measures. For example, therapeutic methods can be used to determine the spread of a B-CLL and lead to its remission.

Therapeutically useful agents include, for example, leukeran, adrimycin, aminoglutethimide, aminopterin, azathioprine, bleomycin sulfate, bulsulfan, carboplatin, carminomycin, carmustine, chlorambucil, cisplatin, cyclophosphamide, cyclosporine, cytarabidine, cytosine arabinoside, cytoxin dacarbazine, dactinomycin, daunomycin, daunorubicin, doxorubicin, esperamicins, etoposide, fluorouracil, ifosfamide, interferon-α, lomustine, melphalan, mercaptopurine, methotrexate, mitomycin C, mitotane, mitoxantrone, procarbazine HCl, taxol, taxotere (docetaxel), teniposide, thioguanine, thiotepa, vinblastine sulfate, vincristine sulfate and vinorelbine. Additional agents include those disclosed in Chapter 52, Antineoplastic Agents (Paul Calabresi and Bruce A. Chabner), and the introduction thereto, pp.1202-1263, of Goodman and Gilman's "The Pharmacological Basis of Therapeutics", Eighth Edition, 1990, McGraw-Hill, Inc. (Health Professions Division). Toxins can be proteins such as, for example, pokeweed anti-viral protein, cholera toxin, pertussis toxin, ricin, gelonin, abrin, diphtheria exotoxin, or Pseudomonas exotoxin. Toxin moieties can also be high energy-emitting radionuclides such as cobalt-60, I-131, I-125, Y-90 and Re-186, and enzymatically active toxins of bacterial, fungal, plant or animal origin, or fragments thereof.

Chemotherapeutic agents can be used to reduce the growth or spread of B-CLL cells and tumors that express the AMB-1 associated epitope of the invention. Animals that can be treated by the chemotherapeutic agents of the invention include humans, non-human primates, cows, horses, pigs, sheep, goats, dogs, cats, rodents and the like. In all embodiments human B-CLL antigens and human individuals are preferred.

Species-dependent antibodies can be used in therapeutic methods. Such a species-dependent antibody has constant regions that are substantially non-immunologically reactive with the chosen species. Such species-dependent antibody is particularly useful for therapy because it gives rise to substantially no immunological reactions. The species-dependent antibody can be of any of the various types of antibodies as defined above, but preferably is mammalian, and more preferably is a humanized or human antibody.

Compositions

Therapeutically useful agents can be formulated into a composition with the antibodies of the invention and need not be directly attached to the antibodies of the invention. However, in some embodiments, therapeutically useful agents are attached to the antibodies of the invention using methods available to one of skill in the art, for example, standard coupling procedures.

Compositions may contain antibodies, antigenic epitopes or trypsin-like protease inhibitors. Such compositions are useful for detecting the AMB-1 protein (for example antigenic epitopes) and for therapeutic methods involving prevention and treatment of B-CLLs associated with the presence of the AMB-1 (for example antigenic epitopes).

The antibodies, (and for example antigenic epitopes and protease inhibitors) can be formulated as pharmaceutical compositions and administered to a mammalian host, such as a human patient in a variety of forms adapted to the chosen route of administration. Routes for administration include, for example, intravenous, intra-arterial, subcutaneous, intramuscular, intraperitoneal and other routes selected by one of skill in the art.

Solutions of the antibodies, (and for example antigenic epitopes and protease inhibitors) can be prepared in water or saline, and optionally mixed with a nontoxic surfactant. Formulations for intravenous or intra-arterial administration may include sterile aqueous solutions that may also contain buffers, liposomes, diluents and other suitable additives.

The pharmaceutical dosage forms suitable for injection or infusion can include sterile aqueous solutions or dispersions comprising the active ingredient that are adapted for administration by encapsulation in liposomes. In all cases, the ultimate dosage form must be sterile, fluid and stable under the conditions of manufacture and storage.

Sterile injectable solutions are prepared by incorporating the antibodies, antigenic epitopes and protease inhibitors in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filter sterilization.

Polynucleotides

In a still further aspect the invention relates to an isolated polynucleotide selected from the group consisting of:

i) a polynucleotide comprising nucleotides of SEQ ID No 5,
ii) a polynucleotide encoding a polypeptide having the amino acid sequence of SEQ ID No 3,
iii) a polynucleotide, the complementary strand of which hybridises, under stringent conditions, with a polynucleotide as defined in any of i) and ii).
iv) a polynucleotide which is degenerate to the polynucleotide of iii), and
v) the complementary strand of any such polynucleotide.

The polypeptides encoded by the polynucleotides may furthermore

a) have at least 60 % sequence identity with the amino acid sequence of SEQ ID No 3 and have interleukin or cytokine activity,

b) be recognised by an antibody, or a binding fragment thereof, which is capable of recognising an epitope, wherein said epitope is comprised within a polypeptide having the amino acid sequence of SEQ ID No 3; and/or

c) be competing with a polypeptide having the amino acid sequence as shown in SEQ ID No 3 for binding to at least one predetermined binding partner such as a cytokine receptor.

Specific examples of fragments of SEQ ID No 1 include the nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18.

Hybridisation

The entire nucleotide sequence of the coding sequence of SEQ ID No 1 and/or SEQ ID No 5 or portions thereof can be used as a probe capable of specifically hybridising to corresponding sequences. To achieve specific hybridisation under a variety of conditions, such probes include sequences that are unique and are preferably at least about 10 nucleotides in length, and most preferably at least about 20 nucleotides in length. Such probes can be used to amplify corresponding sequences from a chosen organism or individual by the well-known process of polymerase chain reaction (PCR) or other amplification techniques. This technique can be used to isolate additional nucleotide sequences from a desired organism or as a diagnostic assay to determine the presence of the coding sequence in an organism or individual. Examples include hybridisation screening of plated DNA libraries (either plaques or colonies; see e. g. Innis et al. (1990) PCR Protocols, A Guide to Methods and Applications, eds., Academic Press).

The terms "stringent conditions" or "stringent hybridisation conditions" include reference to conditions under which a probe will hybridise to its target sequence, to a detectably greater degree than other sequences (e. g., at least twofold over background). Stringent conditions are target sequence dependent and will differ depending on the structure of the polynucleotide. By controlling the stringency of the hybridisation and/or washing conditions, target sequences can be identified which are 100% complementary to a probe (homologous probing).

Alternatively, stringency conditions can be adjusted to allow some mismatching in sequences so that lower degrees of similarity are detected (heterologous probing).

Generally, probes for hybridisation of this type are in a range of about 1000 nucleotides in length to about 250 nucleotides in length.

An extensive guide to the hybridisation of nucleic acids is found in Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes,

Part I, Chapter 2, "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993); and Current Protocols in Molecular Biology, Chapter 2, Ausubel, et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995). See also Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual (2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.).

Specificity is typically the function of post-hybridisation washes, the critical factors being the ionic strength and temperature of the final wash solution.

Generally, stringent wash temperature conditions are selected to be about 5°C to about 2°C lower than the melting point (Tm) for the specific sequence at a defined ionic strength and pH. The melting point, or denaturation, of DNA occurs over a narrow temperature range and represents the disruption of the double helix into its complementary single strands. The process is described by the temperature of the midpoint of transition, Tm, which is also called the melting temperature.

Formulas are available in the art for the determination of melting temperatures.

Preferred hybridisation conditions for the nucleotide sequence of the invention include hybridisation at 42°C in 50% (w/v) formamide, 6X SSC, 0.5% (w/v) SDS, 100 mg/ml salmon sperm DNA. Exemplary low stringency washing conditions include hybridization at 42°C in a solution of 2X SSC, 0.5% (w/v) SDS for 30 minutes and repeating. Exemplary moderate stringency conditions include a wash in 2X SSC, 0.5% (w/v) SDS at 50°C for 30 minutes and repeating.

Exemplary high stringency conditions include a wash in 2X SSC, 0.5% (w/v) SDS, at 65°C for 30 minutes and repeating. Sequences that correspond to the AMB-1 gene or fractions thereof according to the present invention may be obtained using all the above conditions. For purposes of defining the invention, the high stringency conditions are used.

Mutations

Finally, the invention provides a method for determining an increased or decreased predisposition for B-CLL comprising determining in a biological sample from an individual a germline alteration in a target nucleic acid sequence comprising 150,000 nucleotides, said target nucleic acid sequence comprising at least 10 nucleotides of SEQ ID No 1 and/or SEQ ID No:5. This aspect is based on the finding of the importance of the expression product of SEQ ID No 1 and/ SEQ ID No:5, and the complete absence of any detectable expression product of SEQ ID No 1 and/or SEQ ID No:5 in healthy tissue and in patients with good prognosis B-CLL. It is highly likely that the difference is caused by a germline alteration. A germline alteration can be targeted by gene therapy methods and by the methods provided in the present invention.

Preferably, said predisposition is a predisposition for poor prognosis of B-CLL.

Examples

Example 1: cDNA cloning

By Differential Display (Pardee et al., 1992, Jørgensen et al., 1999) part of a gene (hereafter referred to as AMB-1) was found that is expressed in unmutated B-CLL patients with poor prognosis. This gene is not found in the mutated B-CLL patients. When AMB-1 was sequenced and aligned to known sequences in GenBank, perfect homology was found to 225 base pairs (bp) of human genomic DNA from chromosome 12.

RNA was prepared using the RNeasy kit from Qiagen, as described by the manufactor (Qiagen, Hilden, Germany). RNA was prepared from patients with B-CLL without hyper mutation who, by PCR analysis, using primers FDP5 (CCTTTATGTGTGTGACAAGTG) and F10 (ATCCAGCCAGGATGAAATAGAA), showed a high level of the resulting PCR fragment. Poly-A^+ RNA was isolated from total RNA by the "MicroPoly(A)Purist" kit from Ambion, as described by the manufactor (Ambion, Inc., Texas, USA). Cloning-ready cDNA was prepared from 8 µg poly-A^+ RNA using the "ZAP Express® XR Library Construction Kit" from Stratagene as described by the manufactor (Stratagene, San Diego, USA). The cDNA was size fractionated and two size fractions (fraction-1: > 2500 bp and fraction-2 300-2500 bp) were independently ligated to pre-digested lambda Zap vectors and packed into phage particles as described by the manufacture (Stratagene, San Diego, USA). The titer was determined for each library and 200,000 pfu of from the fraction-1 library were plated onto two 22x22 cm screening plates (100,000 pfu on each plate) and 750,000 pfu of the fraction-2 library were plated on five 22x22 cm screening plates (150,000 pfu on each) as described by Stratagene, San Diego, USA. The plates were incubated at 37°C for 18 hours and the plaques transferred to replica nylon filters (Amersham) and denatured and renatured to allow hybridisation. All procedures were made as described by the manufactures (Stratagene, San Diego, USA & Amersham Biosciences, Buckinghamshire, UK).

The filters were screened by independent hybridisations with alpha[^{32}P]-dATP-labelled DNA fragments; alpha[^{32}P]-dATP was purchased from Amersham Biosciences, Buckinghamshire, UK. Between succesive hybridisations, the old probe was removed by incubation for 20 min in 2l 90-100°C water containing 0.1% SDS . The DNA fragments used as probes were (all positions relate to sequence ID # X): 1) pos. 48978-49250; 2) pos. 50011-51591; 3) pos. 51461-52182; 4) pos. 51901-52589; 5) pos. 53121-56521; 6) pos. 58163-59408. All hybridisations and washes were made according to the indtructions from Stratagene, San Diego, USA and Amersham Biosciences, Buckinghamshire, UK; Washing was done at a high stringency (0.1 x SSC at 65°C for 20 min).

A total of 38 plaques that showed a positive response from one or more of the screenings, were excised from the screening plates and grown as plasmids as described (Stratagene, San Diego, USA).

A total of 8 cDNAs were identified by cDNA cloning or by a combination of cDNA cloning, PCR analysis and RACE (rapid amplification fo cDNA ends-polymerase chain reaction) using the the SMART RACE cDNA amplification kit (Clontech, Palo Alto, CA) according to the manufacturer's instructions.

Example 2: Bioinformatic analysis of AMB-1.

The DNA and protein sequence data bases (GenBank and EBI) have been searched for sequences with similarity to AMB1. There is no significant DNA sequence similarity to any known gene. In particular, the coding region of the AMB1 mRNA (SEQ ID no 3) is not present in any known EST. The only significant match to the complete mRNA sequences and the DNA sequence of the putative coding region were BAC clones derived from the region on human chromosome 12 where the gene is located. The "AMB-1 gene" had not been annotated as a gene on the chromosome. Searches with the peptide sequence in the sptrnr data base of peptide sequences (includes Sprot and nrtrembl) showed a low similarity to putative intron maturases from cloroplasts and to bovine IL4. The percentage similarity to both maturases and bovine IL4 was low (25.6% and 30.3%, respectively) and the similarity to maturases only included a match to 75 amino acids of the much larger maturases. In contrast, the match to bovine IL4 extended over the full peptide sequence. IL4, and other 4-helical cytokines, include a leader peptide sequence (signal peptide) allowing the proteins to be secreted. The AMB1 peptide sequence includes a N-terminal peptide sequence with similarity to signal peptide sequences, however, it is not a typical sequence.

A 3D search has been performed, where a peptide sequence is searched for similarity to known protein or peptide 3D-structures. The two best matches were the thioredoxin fold and the human 4-helical cytokine IL4 (Fig. 5). The two matches had almost similar probability scores (2.88 and 3.05, respectively). Searches with 4-helical cytokine peptide sequences (IL4, IL3, IL13 and GM-CSF) revealed that all could be folded into both a 4-helical cytokine structure and the thioredoxin fold. Alignment based on the structural similarity between IL4 (d1iara) and AMB-1 is shown in Fig. 5. Thus, the AMB1 peptide sequence share this property with 4-helical cytokines. The structural similarity is not perfect (Fig. 6) and there are no obvious glycosylation sites in the AMB1 sequence, however, the similarity is significant. Alignment of the AMB1 peptide sequence with the sequences of IL4, IL3, IL13 and GM-CSF, based on their structures, showed very little sequence conservation but a high degree of structural conservation (Fig.7). Based on this alignment, AMB1 has similarities to all the 4-helical cytokines, and the length of AMB1 and the position of gaps in the alignment could suggest a higher similarity to e.g. IL13, but searches at 3D-PSSM only identified a significant similarity to the structure of IL4, not IL13, IL3 or GM-CSF. However, the search algorithms are not perfect and may therefore not detect a possible low structural similarity.

Example 3: Differential expression of AMB-1

Patient material

Blood samples were collected from newly diagnosed untreated patients with B-CLL. Mononuclear cells were isolated by Lymphoprep separation (Nycomed Pharma, Oslo, Norway), and the percentage of CD5+CD20+ B-CLL cells in the mononuclear fraction was >90% in all samples as determined by flow cytometric analysis.

Isolation of RNA and conversion to cDNA.

Material for RNA production was isolated mononuclear cells from B-CLL patients or mononuclear cells from lymphoprep separated buffy coats from normal donors. Total RNA was isolated from $5x10^7$ or more cells using the QIAamp RNA Blood Mini kit (Qiagen, Valencia, CA) with DNAse treatment. RNA (1ug) was converted to cDNA by incubation with a mixture of random-primers (1µg) and T24-primer (1µg) for 5 minutes at 70°C. After cooling on ice, the reaction mixture was added to a final volume of 25µl containing 30U of AMV Reverse Transcriptase HC (Promega, Madison, WI, USA), 1x First Strand Buffer (50mM Tris-HCl, pH 8.3, 50mM KCl, 10mM $MgCl_2$, 10mM DTT, 0.5mM spermidine), 2.5mM of each dNTP and 60U rRNasin ribonuclease inhibitor (Promega, Madison, WI, USA). The reaction was performed for 60 minutes at 37°C.

Determination of somatic hypermutation status

Two µl of cDNA was amplified using a GeneAmp PCR System 2700 (Applied Biosystems, Warrington, UK) with a 40 pmol specific upstream primer corresponding to 1 of the 6 human VH family leader sequences (VH1: 5'-CCATGGACTGGACCTGGAGG-3', VH2: 5'-ATGGACATACTTTGTTCCAGC-3', VH3: 5'-CCATGGAGTTTGGGCTGAGC-3', VH4: 5'-ATGAAACACCTGTGGTTCTT-3', VH5: 5'-ATGGGGTCAACCGCGATCCT-3', VH6: 5'-ATGTCTGTCTCCTTCCTCAT-3') and a 40 pmol downstream primer (Cµ:5'-GAGGCTCAGCGGGAAGACCTT-3' or Cγ:5'-GGGGAAGACCGATGGGCCCCT-3') corresponding to a consensus sequence of the constant region of IgM or IgG respectively. The Reverse Transcription (RT)-PCR reaction contained 1xPCR buffer (10mM Tris-HCl, pH 9.0, 50mM KCl, 0.1% Triton X-100), 2.5mM $MgCl_2$, 0.2mM of each dNTP and 1.5U Taq DNA Polymerase (Promega, Madison, WI, USA) in a final volume of 100µl. The RT-PCR was performed under the following conditions: 1 cycle of 94°C for 5 minutes, 30 cycles of denaturation at 94°C for 30 secs, annealing at 62°C for 30 sec. and extension at 72°C for 30 sec, and a final extension at 72°C for 7 minutes. The RT-PCR products were analysed on 2% agarose gels and sequenced in an HBI Prism 310 Genetic Analyzer (Perkin Elmer, Foster City ,CA, USA) using the BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems, Warrington, UK) following the manufacturer's instructions.

Sequences obtained from each sample were compared to germ line sequences in the V base sequence directory (I.M. Tomlinson, MRC Center for Protein Engineering, Cambridge, UK) using BLAST, and the closest germ line sequence was assigned. A gene sequence was

considered to be mutated if it had equal or more than 2% sequence alterations when compared to the closest published germ line sequence.

RT-PCR that amplifies the Exon 2-Exon 3 junction

To evaluate the mRNA expression pattern of AMB1 in unmutated and mutated B-CLL patients RT-PCR was performed. Exon-overlapping oligonucleotide primers were: 5'-ATCCAGCCAGGATGAAATAGAA-3' and 5'-CACTTGTCACACACATAAAGG-3'. The RT-PCR was performed in a GeneAmp PCR System 2700 thermal cycler with an initial denaturation at 94°C for 2 minutes, 40 cycles of 96°C for 25 sec., 62°C for 25 sec. and 72°C for 90 secs, and a final extension at 72°C for 5 minutes. The reactions contained 2µl cDNA, 1x DDRT-PCR buffer (10mM Tris-HCl, pH 8.3, 50mM KCl, 1.8mM $MgCl_2$, 0.1% Triton X-100, 0.005% gelatine), 0.25mM of each dNTP, 30 pmol of each primer and 0.5U Taq DNA Polymerase (Promega, Madison, WI, USA) in a 30µl final volume. RT-PCR products were analyzed by gelelectrophoresis on 2% agarose gels and visualized with a Gene Genius Bio Imaging System (Syngene, Frederick, MD) after staining with ethidium bromide. An actin control RT-PCR was performed using the primers: 5'-TGACGGGGTCACCCACACTGTGCCCATCTA-3' and 5'-CTAGAAGCATTTGCGGTGGACGATGGAGGG-3'.

RT-PCR that amplifies the Exon 1-Exon 3 junction:

To evaluate the mRNA expression pattern of AMB1 in unmutated and mutated B-CLL patients RT-PCR was performed. Exon-overlapping oligonucleotide primers were: 5'-AGACGGCTCTCACC AATAAG-3' and 5'-CACTTGTCACACACATAAAGG-3'. The RT-PCR was performed in a GeneAmp PCR System 2700 thermal cycler with an initial denaturation at 94°C for 2 minutes, 40 cycles of 96°C for 25 sec., 62°C for 25 sec. and 72°C for 90 secs, and a final extension at 72°C for 5 minutes. The reactions contained 2µl cDNA, 1x DDRT-PCR buffer (10mM Tris-HCl, pH 8.3, 50mM KCl, 1.8mM $MgCl_2$, 0.1% Triton X-100, 0.005% gelatine), 0.25mM of each dNTP, 30 pmol of each primer and 0.5U Taq DNA Polymerase (Promega, Madison, WI, USA) in a 30µl final volume. RT-PCR products were analyzed by gel electrophoresis on 2% agarose gels and visualized with a Gene Genius Bio Imaging System (Syngene, Frederick, MD) after staining with ethidium bromide (Continental Lab Products, San Diego, USA).

Statistical analysis

Statistical significance of the correlation between somatic hypermutation status and AMB1 expression was analyzed using Fisher's exact test.

Northern blotting. RNA from spleen, bone marrow and colon was purchased from Clontech. The AMB1 probe was an 896 base pair fragment (57661-56766) obtained by RT-PCR as described above with the primers 5'-TCACCTGGGAGCTCAGAGGA-3' and 5'-GTGATCCTGGGAGAATCTCT-3'. For Northern blotting, 5 µg of RNA was run on a 1% agarose-gel with 6% formaldehyde dissolved in 1 x MOPS (20 mM 3-(N-morpholino)-propane-sulfonic acid, 5 mM sodium acetate, 1 mM EDTA, pH 7.0) for size separation. The presence of equal amounts of RNA in each lane was ensured by ethidium bromide staining. The RNA was transferred to a Hybond-N membrane (Amersham, Little Chalfont, UK) by

capillary blotting and fixed by UV-irradiation. The filters were pre-hybridized for 1-2 hours at 42°C in 6 ml ULTRAhyb (Ambion, Austin, TX, USA) preheated to 68°C and hybridized overnight at 42°C after addition of further 4 ml containing the ^{32}P-labeled probe and sheared salmon sperm DNA (10 µg/ml). The membranes were washed for 2 x 15 min. at 42°C in 2 x SSC, 0.1% SDS followed by 1 x 15 min. in 0.2 x SSC, 0.1 % SDS and 2 x 15 min. in 0.1 x SSC, 0.1 % SDS at 42°C. The blot was developed and quantified by a phosphoimager. The sizes of the mRNAs were determined by reference to 18S and 28S ribosomal RNA, which were visualized by ethidium bromide staining. The AMB1 probe used for hybridization was radiolabeled with [α-^{32}P] dCTP using the Random Primers DNA Labeling System (Gibco BRL).

Dot blot of multiple tissue expression (MTE) array. An MTE array (Clontech, Palo Alto, CA, USA) was hybridised to AMB1 at 65°C in ExpressHyb (Clontech) supplemented with sheared salmon sperm DNA (7.5 µg/ml) and human C_0t-1 DNA (1.5 µg/ml) according to the manufacturers recommendations. The tissue types represented on the MTE array are shown in Figure 11. Following hybridisation the filter was washed 5 x 20 min. at 65°C in 2 x SSC (1 x SSC =150 mM NaCl, 15 mM sodium citrate, pH 7.0), 1% SDS and 2 x 20 min at 65°C in 0.1 x SSC, 0.5% SDS. The blot was developed and quantified by a phosphoimager (Fuji Imager Analyzer BAS-2500, Image Reader ver. 1.4E, Image Gauge ver. 3.01 software, Fuji, Stockholm, Sweden). The membranes were stripped by boiling in 0.5% SDS for 10 min. before rehybridization. The probe used for hybridization were radiolabeled with [α-^{32}P] dCTP using the Random Primers DNA Labeling System (Gibco BRL, Rockville, ML, USA).

Results

Based on the known sequence of the AMB-1 cDNA RT-PCRs with primers that extend across the Exon 2-Exon 3 junction and the Exon 1- Exon 3 junction were set up. As shown in Figure 10, where the Exon 2- Exon 3 junction has been amplified, AMB-1 is expressed in the unmutated patients (UPN1-8) while no expression of AMB-1 is seen in mutated patients (UPN9-16).

Northern blot analysis was performed to determine the size of AMB-1's mRNA transcript. As shown in Figure 3 the probe identifies transcripts in the samples from the three patients without somatic hypermutation (UPN1, UPN4 and UPN7). However, the probe does not recognise any transcripts from the patients with somatic hypermutation (UPN9, UPN10, UPN13, UPN21) or the various cell lines and tissue samples. Similar results were obtained when cell lines and tissue samples were investigated by RT-PCR (results not shown). Dot blot analysis on a purchased filter with 96 different RNA samples (Figure 8) only revealed specific binding to the total DNA control dot, but not to any specific tissue (results not shown). Thus AMB-1 is only expressed in B-CLL cells without hypermutation or AMB-1 is expressed at extremely low levels in other tissues.

We next tested the predictive value, in terms of IgV_H mutational status, of expression of AMB-1 in 29 consecutive newly diagnosed patients. At present 13 somatically unmutated and 16 somatically mutated patients have been included in our prospective patient database. The sensitivity and specificity for expression of AMB-1 in predicting mutational status is well above 90% ($p<0.0001$), which is at the level obtained by sequencing.

Example 4. Investigation of the prognostic significance of AMB-1 in terms of patient survival

To obtain information about the ability of AMB-1 to predict survival or time to progression for B-CLL patients survival curves were made. For each patient the following data were collected: clinical stage at time of diagnosis (Rai and Binet staging), date of diagnosis, date of first time of treatment and last follow up date. Time to treatment (progression free survival) and survival time were calculated based on these dates using the Microsoft Excel software. Survival times and progression free survival times were plotted be the Kaplan-Meier Method and compared using the log-rank test (Prism 3.0 Graph Pad software).

34 newly diagnosed, untreated B-CLL patients were investigated as described above. As shown in Figure 2a-d AMB-1 expression is a good predictor of B-CLL patient survival and B-CLL patient time to progression.

Example 5: siRNA assays.

Materials and Methods:

Frozen cells from unmutated or mutated B-CLL patients pruified and characterized for mutational status as previously described were thawed and Ficoll separated to obtain the live cells. Cells were counted and resuspended at $2x10^6$ cells/ml in RPMI 1640 medium with glutamax-1 (RPMI, GIBCO, Paisley, UK). $1x10^6$ cells (500µl of cell suspension) were incubated on ice for 10 minutes with the appropriate siRNA or combination of siRNAs at a final concentration of 100nM each. The suspension was electroporated using a 4mm cuvette (Molecular BioProducts, San Diego, CA, USA) in a BioRad Gene Pulser. Following electroporation the cells were incubated on ice for 10 minutes, washed once in RPMI with 20% FCS and 100units/ml Penicillin and 100µg/ml Streptomycin (GIBCO, Paisley, UK), resuspended at a concentration of $4x10^6$ cells/ml in the same medium and incubated for 48 hours. For each patient the optimal voltage allowing for maximum siRNA uptake was determined by making a voltage curve (300-450V, 960µF) using a FITC-labeled unspecific probe (Xeragon, Qiagen, Hilden, Germany) at a final concentration of 200nM and the uptake was followed by flow cytometry.

The siRNAs were produced using the SilencerTM siRNA Construction Kit, (Ambion, Texas, USA) according to the manufacturers instructions. SiRNAs had the following target sequences:

siRNA name:	mRNA target sequence:	

Exon 1 A	5'- AAUAAGGGCAGGCAUCAUCCA -3'	SEQ ID No:19
Exon 1 B	5'- AAUUACACUGCCAGGUUUCCU -3'	SEQ ID No:20
Exon 2 A	5'- AAUUCAUUCACAAUGAUUGCU -3	SEQ ID No:21
Exon 2 B	5'- AAUUUCUCUUGGGUAAUUCAG-3'	SEQ ID No:22
Exon 3 (DDend) A	5'- AAAAUCAGAAUCUGCGCAGCA -3'	SEQ ID No:23
Exon 3 (DDend) B	5'- AAUGAUGAUGGGAAGAAGGAA -3'	SEQ ID No:24
CDS A	5'- AAACUUAGUAAUUGAGUGUGA 3'	SEQ ID No:25
CDS B	5'- AAUAUGUCACUUUCAUAAAGC -3'	SEQ ID No:26
Transcript on -strand A	5'- AAUGAUGAUGGGAAGAAGGAA -3'	
Transcript on -strand B	5'- AAACUAUGAGAUUUCAGAAGG -3'	

Table 1: Overview of the siRNAs used in electroporation experiments

After 48 hours of incubation the various B-CLL samples were counted and live and dead cells were distinguished by nigrosin exclusion (0.1% in Nigrosin in PBS from Fluka, Buchs, Switzerland).

Results:

Following 48 hours of incubation with control GFP siRNA or siRNAs against various regions of the cDNAs (see table 1), the viable cells and dead cells were counted based on the ability of the cells to exclude nigrosin. As shown in table 2, in 3 out of 4 patients introduction of siRNAs against Exon 1, Exon 2 and Exon 3 resulted in an increase of dead cells as compared to the controls (no siRNA or GFP siRNA).

Treatment:	UPN 67	UPN 66	UPN 62	UPN 73
No siRNA	33.0%	19.7%	30.9%	20.2%
Control siRNA	28.4%	26.7%	24.0%	16.0%
CDS siRNA	23.7%	25.5%	21.0%	13.3%
Ex1-3 siRNA	47.3%	54.1%	17.2%	29.1%

Table 2: Percent dead cells of total cells following electroporation and incubation with siRNAs.

Example 6: Identification of possible cytogenetic aberrations near or within the region encoding AMB-1 on chromosome 12.

Rationale: The limited expression profile of AMB-1 suggests that it may be a result of a genetic aberration (e.g. deletion, translocation or alternative splicing) or that the promotor region controlling the expression of AMB-1 is uniquely activated in unmutated B-CLL. Another gene is situated about 200.000 bases upstream of the AMB-1 gene (SEQ ID No 1) on chromosome 12 and the inventors we have determined that this gene is expressed at equal levels in unmutated and mutated patients.

Methods: Using primers, initially spaced about 20.000 bp apart; this region on chromosome 12 is characterised in unmutated B-CLL patients. If genetic aberrations within the region are detected by PCR analysis of chromosomal DNA, detailed molecular genetic studies using FISH, microsatellite analysis and Southern blotting will be employed. The whole region from unmutated patients is sequenced.

Example 7: Polyclonal antibodies

Production of polyclonal antibodies:

Synthetic peptides CDLETNSEINKLIIYLFSQNNRIRF and CQVSKKHIIYSTFLSKNF were synthesized and conjugated to KLH (K.J.Ross-Petersen Aps, Holte, Denmark). Polyclonal antibodies were produced by immunization of rabbits with these conjugated peptides by DAKO (DAKO Cytomation A/S, Glostrup, Denmark).

Testing of polyclonal antibodies:
At the present time we have produced polyclonal antibodies from three rabbits that have been immunized with peptides representing predicted immunogenic regions of the protein that can be predicted from the CDS sequence (SEQ ID No:17) (cDNA 4). The antibodies are tested in various ways. The proposed reading frame of CDS (SEQ ID No:17) is expressed in 293 cells and the binding of antibodies to lanes on a western blot with non-transfected 293 cells versus transfected 293 cells are compared. The size of the band in the lanes with transfected 293 cells is compared to the size of western blot bands in lanes with proteins from B-CLL patients. Specificity of the bands is secured by peptide blocking experiments.

Additionally the polyclonal antibodies are tested in B-CLL immunoprecipitation experiments where the antibodies are used to immunoprecipitate the protein produced from the CDS sequence and the immunoprecipitates are analyzed by western blotting.

Example 8: Assay for the biological activity of 4-helix cytokines.

The assay is based on the use of a cytokine dependent or stimulated cell line, for example an IL4 dependent cell line ("Optimisation of the CT.h4S bioassay for detection of human interleukin-4 secreted by mononuclear cells stimulated by phytohaemaglutinin or by human leukocyte antigen mismatched mixed lymphocyte culture", Petersen, S.L., Russell, C.A., Bendtzen, K. & Vindeløv, L.L., Immunology Letters 84 (2002) 29-39). Other examples of cytokine dependent cell lines include IL13 dependent cell lines. A list of commercially available cytokine dependent cell lines is disclosed in the general part of the description. These can all be used for assessing cytokine activity. The most preferred cell lines are those that are IL4 dependent.

The assay can be performed in two ways. The first assay comprises providing recombinantly produced AMB1 protein or a functional equivalent thereof and determine the proliferation rate of the cell line. The proliferation rate (either rate of proliferation or ± proliferation) can be compared to the proliferation rate of the cell line exposed to IL4 or another known 4-helical cytokine or interleukin.

If a positive result is obtained with a polypeptide an assay will be performed on the same cell line with the IL4 receptor blocked. This will check whether the stimulus goes through IL4R.

The second assay is based on transfection of a gene encoding a 4-helix cytokine according to the invention into cytokine dependent cells and observe proliferation or non-proliferation during transient expression.

Example 9: Cytokine receptor binding assays

The following is a description of the layout of a cytokine receptor binding assay used to determine the cytokine activity of the 4-helix cytokines according to the present invention.

The assays can be performed with any cytokine receptor. Preferred receptors include but is not limited to the receptors for IL4 IL13, IL3, and GM-CSF.

The ability of recombinant cytokine receptor to bind to 4-helical cytokine is assessed in a competitive binding ELISA assay as follows. Purified recombinant cytokine receptor (IL4, IL13, IL3 or GM-CSF receptors) (20 µg/ml in PBS) is bound to a Costar EIA/RIA 96 well microtiter dish (Costar Corp, Cambridge Mass., USA) in 50 µL overnight at room temperature. The wells are washed three times with 200 µL of PBS and the unbound sites blocked by the addition of 1% BSA in PBS (200 µl/well) for 1 hour at room temperature. The wells are washed as above. Biotinylated AMB-1 (1 µg/ml serially diluted in twofold steps to 15.6 ng/mL; 50 µL) is added to each well and incubated for 2.5 hours at room temperature. The wells are washed as above. The bound biotinylated AMB-1 is detected by the addition of 50 µl/well of a 1:2000 dilution of streptavidin-HRP (Pierce Chemical Co.,

Rockford, Ill.) for 30 minutes at room temperature. The wells are washed as above and 50 µL of ABTS (Zymed, Calif.) added and the developing blue color monitored at 405 nm after 30 min. The ability of unlabelled 4-helical cytokine to compete with biotinylated AMB-1, respectively, is assessed by mixing varying amounts of the competing protein with a quantity of biotinylated AMB-1 shown to be non-saturating (i.e., 70 ng/mL; 1.5 nM) and performing the binding assays as described above. A reduction in the signal (Abs 405 nm) expected for biotinylated 4-helical cytokine indicates a competition for binding to immobilised cytokine receptor.

The above identified assays can be used to identify 4-helical cytokines with similar binding affinities as AMB-1 (SEQ ID No. 3). In the competitive binding assays biotinylated IL4, IL13, IL3, or GM-CSF can be used to identify 4-helical cytokines which can compete with these cytokines.

References

1. U.S. Pat. No. 4,554,101
2. U.S. Pat. No. 5,681,729
3. Antineoplastic Agents (Paul Calabresi and Bruce A. Chabner
4. Andrews, R., L. Rosa, M. Daines, and G. Khurana Hershey. 2001. Reconstitution of a functional human type II IL4/IL13 receptor in mouse B cells: demonstration of species specificity. *J Immunol* 166, no. 3:1716
5. Ausubel, et al., Current Protocols in Molecular Biology, Chapter 2, Eds., Greene Publishing and Wiley-Interscience, New York (1995).
6. Bouteiller, C.L., R. Astruc, A. Minty, P. Ferrara, and J.H. Lupker. 1995. Isolation of an IL13-dependent subclone of the B9 cell line useful for the estimation of human IL13 bioactivity. *J Immunol Methods* 181, no. 1:29
7. Chaouchi, N., C. Wallon, C. Goujard, G. Tertian, A. Rudent, D. Caput, P. Ferrera, A. Minty, A. Vazquez, and J.F. Delfraissy. 1996. Interleukin-13 inhibits interleukin-2-induced proliferation and protects chronic lymphocytic leukemia B cells from in vitro apoptosis. *Blood* 87, no. 3:1022
8. Damle, R.N., T. Wasil, F. Fais, F. Ghiotto, A. Valetto, S.L. Allen, A. Buchbinder, D. Budman, K. Dittmar, J. Kolitz, S.M. Lichtman, P. Schulman, V.P. Vinciguerra, K.R. Rai, M. Ferrarini, and N. Chiorazzi. 1999. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood* 94, no. 6:1840
9. Dancescu, M., M. Rubio-Trujillo, G. Biron, D. Bron, G. Delespesse, and M. Sarfati. 1992. Interleukin 4 protects chronic lymphocytic leukemic B cells from death by apoptosis and upregulates Bcl-2 expression. *J Exp Med* 176, no. 5:1319
10. Dohner, H., S. Stilgenbauer, A. Benner, E. Leupolt, A. Krober, L. Bullinger, K. Dohner, M. Bentz, and P. Lichter. 2000. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med* 343, no. 26:1910
11. Fluckiger, A.C., F. Briere, G. Zurawski, J.M. Bridon, and J. Banchereau. 1994. IL13 has only a subset of IL4-like activities on B chronic lymphocytic leukaemia cells. *Immunology* 83, no. 3:397

12. Goodman and Gilman's "The Pharmacological Basis of Therapeutics", Eighth Edition, 1990, McGraw-Hill, Inc. (Health Professions Division)
13. Hamblin, T.J., Z. Davis, A. Gardiner, D.G. Oscier, and F.K. Stevenson. 1999. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood* 94, no. 6:1848
14. Innis et al. (1990) PCR Protocols, A Guide to Methods and Applications, eds., Academic Press
15. Jørgensen M, Bévort M, Kledal TS, Hansen BV, Dalgaard M, and Leffers H (1999) Differential Display Competitive PCR: An Optimal Tool for Assaying Gene Expression. Electrophoresis 20:230-240
16. Kirn D. et al., "The emerging fields of suicide gene therapy and virotherapy", Trends Mol. Med. 2002, 8:568-573
17. Kröber, A., T. Seiler, A. Benner, L. Bullinger, E. Bruckle, P. Lichter, H. Dohner, and S. Stilgenbauer. 2002. V(H) mutation status, CD38 expression level, genomic aberrations, and survival in chronic lymphocytic leukemia. *Blood* 100, no. 4:1410
18. Kyte J and Doolittle RF. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, no. 1:105
19. Liang P, and Pardee AB (1992) Differential Display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257:967-71
20. Lin, K., P.D. Sherrington, M. Dennis, Z. Matrai, J.C. Cawley, and A.R. Pettitt. 2002. Relationship between p53 dysfunction, CD38 expression, and IgV(H) mutation in chronic lymphocytic leukemia. *Blood* 100, no. 4:1404
21. Lundin, J., E. Kimby, L. Bergmann, T. Karakas, H. Mellstedt, and A. Osterborg. 2001. Interleukin 4 therapy for patients with chronic lymphocytic leukaemia: a phase I/II study. *Br J Haematol* 112, no. 1:155
22. Luo, H.Y., M. Rubio, G. Biron, G. Delespesse, and M. Sarfati. 1991. Antiproliferative effect of interleukin-4 in B chronic lymphocytic leukemia. *J Immunother* 10, no. 6:418
23. Merrifield, J. Am. Chem. Soc. 85:2149-2146, 1963)
24. Oscier, D.G., A.C. Gardiner, S.J. Mould, S. Glide, Z.A. Davis, R.E. Ibbotson, M.M. Corcoran, R.M. Chapman, P.W. Thomas, J.A. Copplestone, J.A. Orchard, and T.J. Hamblin. 2002, Multivariate analysis of prognostic factors in CLL: clinical stage, IGVH gene mutational status, and loss or mutation of the p53 gene are independent prognostic factors. *Blood* 100, no. 4:1177
25. Panayiotidis, P., K. Ganeshaguru, S.A. Jabbar, and A.V. Hoffbrand. 1993. Interleukin-4 inhibits apoptotic cell death and loss of the bcl-2 protein in B-chronic lymphocytic leukaemia cells in vitro. *Br J Haematol* 85, no. 3:439
26. Petersen, S.L., Russell, C.A., Bendtzen, K. & Vindeløv, L.L., Optimisation of the CT.h4S bioassay for detection of human interleukin-4 secreted by mononuclear cells stimulated by phytohaemaglutinin or by human leukocyte antigen mismatched mixed lymphocyte culture", *Immunology Letters* 84 (2002) 29-39)
27. Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual (2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.).

28. Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes, Part I, Chapter 2, "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993)
29. Yazawa et al., Current Progress in suicide gene therapy for cancer, World J. Surg. 2002, 26(7):783-789.